AFOSR 69-1405TR

Development of a Taxonomy of Human Performance:

A Review of Classificatory Systems Relating to Tasks and Performance

1. This document has been approved for public release and sale; its distribution is unlimited.

George R. Wheaton

Technical Report 1

DECEMBER 1968

AMERICAN INSTITUTES FOR RESEARCH
WASHINGTON OFFICE
Address: 8555 Sixteenth Street, Silver Spring, Maryland 20910
Telephone: (301) 587-8201

R68-7

56

AMERICAN INSTITUTES FOR RESEARCH

WASHINGTON OFFICE

Edwin A. Fleishman, Ph.D., Director

INSTITUTE FOR RESEARCH ON HUMAN PERFORMANCE

George H. Johnson, Ph.D., Director
Robert G. Kinkade, Ph.D., Associate Director
Harold P. Van Cott, Ph.D., Research Advisor

Research on human learning and performance as they relate to the functioning and effectiveness of social, educational, and technological systems.

COMMUNICATION RESEARCH PROGRAM

Arthur L. Korotkin, Ph.D., Director
Christopher L. Faegre, Ed.M., Assistant Director

Research on human information processing, communication, and psycholinguistics and on the design and evaluation of communication and educational systems.

SKILLS RESEARCH PROGRAM

Robert G. Kinkade, Ph.D., Director
Armand N. Chambers, Ph.D., Assistant Director

Research on the identification of perceptual, motor, and cognitive skills and the factors affecting their acquisition, maintenance, and retention.

INSTITUTE FOR RESEARCH ON ORGANIZATIONAL BEHAVIOR

Albert S. Glickman, Ph.D., Director

Research on individual, interpersonal and group behavior as they relate to organizational functioning and effectiveness.

HUMAN RESOURCES RESEARCH PROGRAM

Clifford P. Hahn, M.S., Director
Dorothy S. Edwards, Ph.D., Assistant Director

Studies on personnel selection, job proficiency, training procedures, motivation and job satisfaction. A.I.R.'s ACCIDENT RESEARCH CENTER is administered within this program.

MANAGEMENT RESEARCH PROGRAM

Albert S. Glickman, Ph.D., Director

Research on the administrative process and on the social and organizational factors which enhance individual and institutional competence.

INTERNATIONAL RESEARCH INSTITUTE

Paul Spector, Ph.D., Director
Henry P. David, Ph.D., Associate Director
Stanley Lichtenstein, Ph.D., Director of Studies

Research on the development of human resources in developing countries, problems of working effectively abroad, evaluation of action programs in the underdeveloped countries, role of attitudes and values in social change and economic development.

AIR-726-12/68-TR-1

DEVELOPMENT OF A TAXONOMY OF HUMAN PERFORMANCE:

A Review of Classificatory Systems Relating to Tasks and Performance

George R. Wheaton

TECHNICAL REPORT NUMBER 1

Prepared under Contract for
Advanced Research Projects Agency
Department of Defense
ARPA Order No. 1032

Principal Investigator: Edwin A. Fleishman

Contract No. F 44620-67-C-0116

American Institutes for Research
Washington Office
Skills Research Program

December 1968

ABSTRACT

In recent years there has been a cresive interest in the science of taxonomy in general, and in behavioral classifications in particular. In this report, past efforts are reviewed in an attempt to assess the "state-of-the-art" and to provide procedural guidelines for future taxonomic efforts. Approaches to and dilemmas encountered in attempting to develop systems of classification are discussed. Special consideration is given to those systems oriented toward the organization and understanding of information about human task performance. Within this context the taxonomic issues of purpose, descriptive bases, and methodological approaches are discussed in terms of available alternatives. The report leads to the conclusion that behavioral taxonomy is still in its infancy and that truly powerful systems of classification have yet to be developed. The paper suggests how substantive progress may be made by attempting development of a task classification system based upon numerical taxonomic procedures.

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

One of the major problems confronting behavioral scientists and technologists is the lack of a generally accepted set of unifying concepts for the systematic description, prediction, and manipulation of human behavior. Lacking an organizing framework, experimentalist and technologist alike find generalization, communication, and application of research findings to be difficult. The behavioral scientist sifts through seemingly unrelated masses of data in search of even a rudimentary theory of behavior. The behavioral technologist struggles with the application of a seemingly unlimited number of principles. Neither the scientist nor the technologist is able to consistently and systematically relate his results to those from previous studies, to "similar" situations with which he has yet to deal, or to the findings of researchers and technologists working on allied problems. This state-of-affairs argues strongly for a mechanism to provide the needed structuring. More specifically, psychologists must provide organization by coming to grips with the complex taxonomic problems of their discipline.

Although many of these issues are as old as psychology itself, they have been especially disconcerting since World War II, the time when engineering psychology began to emerge as a discipline (Grether, 1968). The engineering psychologist viewed man's performance on a task within a system context. Means were sought to optimize performance through function allocations, design, selection, and training. As new systems developed, vast amounts of data were generated by each of these activities. These experiences and data, however, provided few rigorous procedural guidelines for the development of subsequent systems. Systematic use of principles, and rules for their application were essentially lacking. Neither tasks nor systems could be compared and contrasted on any consistent basis. This situation prompted Melton and

Briggs (1960) to write in the Annual Review of Psychology:

> "It is clear to those working in the area of engineering psychology, and it should become clear to others, that this vigorous and expanding universe of knowledge has semantic and taxonomic problems which have not been overcome. Nor can they be overcome in any stable way by the ingenuity of the organizers of its literature. The roots of these difficulties are many, not the least being the semantic and taxonomic problems of experimental psychology ... Foremost among deficiencies of this type is the lack of taxonomies of tasks or of skills" (p. 89).

Paul Fitts (1962) writing in the book "Training Research and Education" stated the same issue in a learning context:

> "The importance of an adequate taxonomy for skilled tasks is widely recognized in all areas of psychological theorizing today. A taxonomy should identify important correlates of learning rate, performance level, and individual differences. It should be equally applicable to laboratory tasks and to tasks encountered in industry and in military service" (p. 178).

More recently, experimental psychologists have expressed a similar concern. Among others, Fleishman (1967a, 1967b) and Hackman (1968) have stressed the need for a taxonomy of human performance which would allow for the integration of laboratory research and the generalization of such research to operational settings. In spite of these expressions of concern, however, a sophisticated taxonomy of human performance has yet to be developed. We still lack a comprehensive system for the description and classification of tasks performed by humans.

Why, in light of the obvious importance of this problem, has there not been more progress? Why hasn't there evolved a widely accepted and useful classification of tasks or of performance? It can be argued that our relatively slow progress is a function of at least three complex taxonomic issues. In essence they represent choices which must be made about the following:

1. objectives to be served by any particular system of classification;

2. descriptive bases upon which a system of classification is to be founded;

3. methodologies and analytical techniques employed to establish and validate a classification system.

It is the intent of the present report to explore these and related issues in some depth as an aid to future task or performance classification efforts. Toward this end we have reviewed relevant literature bearing on previous taxonomic approaches and concepts in the behavioral sciences. The report describes these efforts briefly, and attempts to synthesize the issues which emerge in dealing with this complex and critical problem.

Scope of the Review

Anticipating the knowledgeable reader who expects an exhaustive survey, let us stress that our review of the literature has been intentionally restricted in scope. Primary emphasis has been placed on literature dealing with the classification of different aspects of human performance and of tasks. Even within this structure exhaustive treatment has not been attempted. Rather, the most recent, comprehensive, or representative efforts of individuals concerned with the problem have been examined. In reviewing these efforts, the intent has been to acquire

and synthesize information on different classification schemes which have been proposed, the purposes to be served by each, and the approaches which have been suggested or employed for their development. We are less interested in describing the particulars of each system than we are in assessing the "state-of-the-art" and providing guidelines for future taxonomic efforts in the area of human task performance.

Approximately 50 studies and papers were selected as being particularly germane to the classification of human performance, behavior, or tasks. Not all of these represent formal classificatory studies. Indeed, the majority discuss classification in rather general terms, while others emphasize a completely descriptive approach -- task analysis. Because of the diversity of this literature synthesis is difficult; but examination of the range of opinions may prove fruitful.

CLASSIFICATORY SYSTEMS AND ISSUES

Classification refers both to a process and to a resultant product. On the one hand, classification is the act or process of systematically arranging some subject matter into groups or categories according to selected criteria. On the other hand, classification is the more or less formally structured set of classes or categories which emerges. Behavioral scientists attempting to organize information about tasks and performance should obviously be interested in classification both as process and as product. Unfortunately, however, this dual interest is seldom expressed. Emphasis is usually placed upon a discussion of alternative structures (products) rather than upon the systematic examination of the general principles and issues of the classification process. It is toward this area that we, as psychologists, must turn our attention.

The penchant to study the product of classification at the expense of the process is well known, having beset and handicapped other disciplines as well. It tends to be characterized on one hand by rather fragmentary and isolated attempts at classification which defy generalization, and on the other by schemas which seldom progress beyond the conceptual level of development. With but a few exceptions, the literature which was sampled for the present report reflected this state-of-affairs.

Basic to the consideration of classification as a process are three major taxonomic issues. The first of these involves the purpose or objective which gives rise to or motivates the desire to classify. Why does the behavioral scientist attempt classification? What does he expect to accomplish from such a complicated and time-consuming

activity? The second issue confronting the behavioral scientist is two-fold. Having stated his purpose in undertaking classification, he must choose appropriate subject-matter and must conceive of ways in which it can be most clearly and systematically described. Is he to classify tasks, performance, or some other subject-matter? In any event, what will his descriptors be based upon -- observed behaviors, hypothesized intervening processes, required abilities, response measures, display-control characteristics, etc.? Last, but certainly not least, the third or methodological issue must be faced. A method of classification must be developed. Criteria for class inclusion or exclusion must be formulated. The applicability of various analytical techniques must be determined. Only when these issues have been squarely faced and resolved, can classification proceed on a logical and consistent basis. Let us now turn to consideration of these issues.

Purpose of Classification

Why are behavioral scientists interested in classification? The question is important because individuals who attempt classification usually do not view the development of such a system, in and of itself, as an end. Rather, they view a system of classification as a tool which is to be employed on behalf of some other goal. In nearly every instance, their goal is the increased ability to interpret, predict, or control some facet of performance (Cotterman, 1959). This goal is to be achieved by seeking relationships between that which is classified (e.g., tasks, processes mediating performance, etc.) and selected variables of interest to a particular investigator (e.g., distribution of practice, training regimens, environmental stressors, etc.).

Desiring to establish such relationships, however, the behavioral scientist must initially decide upon one of two objectives. He can elect to develop a system of classification having utility for a very specific

and, consequently, limited area (e.g., the classification of tasks with respect to which of ten training regimens is most efficacious in promoting high levels of task performance). In this case the objective is oriented toward a particular problem area and is to be achieved by developing an unique classificatory system having little utility for other problem areas. Specific utility is perhaps maximized but is done so at the expense of classificatory generalizability. For example, a specific classificatory system designed to be of aid in the interpretation, prediction, or control of training phenomena may contribute relatively little to the interpretation, prediction, or control of environmental stressor phenomena. (Nor is it intended to do so.)

There is, however, a second type of objective toward which the behavioral scientist can orient his system of classification. He can opt for a system from which a variety of applications may stem. He can develop a classification of tasks, processes, or some other concept and <u>then</u> attempt to organize a wide range of data in terms of his system. For example, he might first classify tasks and <u>only then</u> relate stressors, learning principles, training regimens, etc. to each class of tasks in his system. In this case, classification is designed from inception to be general. It is designed to serve a variety of users by aiding in the interpretation, prediction, or control of a broad range of phenomena.

Implied by the above distinction between objectives is the notion that classifications may be viewed either as means toward a specific end, in which case they are clearly utilitarian, or they may be viewed as ends in and of themselves with eventual application being an essential but totally independent problem. In this latter instance classification may be seen as a step in the development of theory. Any and all tasks, for example, may be systematically described, interrelated, and categorized in terms of their <u>intrinsic</u> properties. Being able to describe tasks in terms of one <u>common</u> and <u>consistent</u> system, it may then be possible to

systematically describe, interrelate and, eventually, explain facets of performance stemming from the different types or categories of tasks.

While the distinction between utilitarian and theoretical classification is useful for our discussion, it must be realized that it is somewhat arbitrary. Any classification is likely to have both applied and theoretical implications. The difference is one of emphasis.

Utilitarian Classifications with Specific Applications

It is not surprising that classifications have been suggested or developed with a variety of applications in mind. Many of these systems are focused on specific areas of content, particularly when they are to be utilized in the solution of human-factors, human-engineering problems. Of the classification systems having rather specific applied objectives, those dealing with training are most numerous. A number of investigators (Annett & Duncan, 1967; Cotterman, 1959a, 1959b; Eckstrand, 1964; Folley, 1964a, 1964b; Gagne, 1962; E. E. Miller, 1966; and Stolurow, 1964) have been concerned with the systematic application of learning principles and training methods to specific types of tasks. As a consequence they have prop sed systems which will permit the classification of tasks* into sets or categories which are relatively homogeneous and invariant with respect to principles of learning, training techniques, etc. The objective of such systems would be to supply training personnel with an explicit rationale for the selection of specific training programs for specific tasks. Bloom (1967) has attempted to develop a similar system for the educational community. By developing a classification of educational objectives, he hopes to facilitate communication among its pedagogical users about appropriate

* Task in this instance is employed as a generic concept including its physical attributes, human functions, performance requirements, etc.

methods of instruction.

Dealing with an important human-engineering problem, Alexander and Cooperband (1965) and Kidd (1962) have considered techniques to provide for effective design in complex systems. They would ultimately wish for categories of tasks which remain relatively invariant for principles of design.

Other investigators (Fitts, 1962; R. B. Miller, 1966) have proposed more comprehensive classification systems, in the sense that they would provide for more than one specific application. Fitts viewed the objective of "task" classification to be the identification and application of important correlates of learning rate, performance level, and individual differences. Miller discussed classification systems which would permit the specification of selection and training requirements, of types of error to be expected, and of design requirements. McCormick (1964) discussed the classification of jobs with similar applications in mind including predictions of job-success.

Theoretical Classifications with Broad Applications

Classificatory systems may also be developed, however, as autonomous structures which are only some time later to be related to other variables or "applied." In these instances, the classificatory exercise is an integral and inextricable step in the development of theory. The resultant system provides a consistent conceptual framework, the elements of which eventually are to be utilized in the interpretation, prediction, or control of behavioral phenomena. One is not precluded from seeking specific applications for such classifications. The point is that he must not let a specific application dictate the composition and structure of his system.

Learning theorists have been most prolific in these pursuits (Melton, 1964). For years, the question of types of human learning has provided a heated controversy -- one which is likely to continue for some time. As would be expected, the longstanding nature of this issue has led to the generation of a number of classifications and to their constant revision. As yet, no categorization has been proposed which effectively compares, contrasts, and interrelates the various "categories." Consequently, we have been unable to formulate a general theory of learning. We shall see shortly, however, that the concept of task and its classification may have ramifications for this area (Wickens, 1964).

The ability theorists have been engaged in analogous work. They have attempted to isolate basic dimensions of behavior upon which a general theory of human performance might be based. Many of these investigators (Fleishman, Guilford, Thurstone, and Cattell), working with factor analytic methods applied to test measures, might not consider themselves as taxonomists, but indeed they are. They differ from the learning theorists in their analysis of relationships among responses to tests rather than among the tests or tasks performed. Fleishman (1964) and Guilford (1967) have been most explicit in attempting to integrate the ability dimensions identified within the general framework of experimental psychology. In particular, Fleishman and his associates have conducted studies relating ability dimensions isolated in perceptual-motor research to stages of learning (1955), stimulus-response relations (1956), retention (1962), part-whole task relations (1965), effects of drugs (1966), etc. This variety of studies was possible because Fleishman first attempted to develop a standard and consistent classificatory structure of human abilities. Eventual integration of these and similar studies is feasible. This feasibility is engendered by the consistent conceptual framework or classificatory system prevading all of the studies.

Other theoretical classifications are being developed as well and are no less important than the efforts previously considered. They do, however, represent different content areas. Bergum (1966) is working on a classification of performance tasks as part of a more general physiologically based theory of performance. Similarly, Hackman (1968) discusses the need for a classification of tasks per se in order to then permit more systematic research on stress.

Summary of Classificatory Objectives

Of importance for future task taxonomic efforts is a distinction among alternative classification systems in terms of their objectives. Two general types of objectives can be identified. One can attempt to relate the system to a specific content area for a specific application or he can relate it to a broad range of content areas. When a specific application is intended, it often dictates the classificatory structure from the start. The general approach seems to be one of grouping "tasks" as a function of the effects of a selected set of variables on measures of task performance. Consequently, grouping of tasks can be achieved regardless of their intrinsic similarities and dissimilarities. On the other hand, in developing classification systems designed to satisfy a much broader range of applications, the approach is altogether different. Direct interest initially lies not in the similarity of effects upon task performance, but rather in the similarity of characteristics of the tasks themselves. This distinction is rarely made in present research practice.

An implication of this distinction must be considered. When a specific application dictates classification, an unique system will be required. For each new content area, a different classification will be necessary. Establishing a function to effectively translate one system of classification into another, in the attempt to synthesize, interrelate,

and integrate data, would be difficult if not impossible. Conversely, one might elect to develop a classification based upon task dimensions, properties, characteristics, etc., without initial concern for area of application. Applications could, however, be eventually sought by relating other sets of variables (for instance, training methods) to the task dimensional values defining a specific task group. Employing a consistent system of classification, communication among different tasks and variables is direct. A translation service is not needed.

The choice of objectives is ours. Having made that choice, however, we may be constrained in terms of our definitions of "task," the unit characteristics we employ and the analytical approaches we pursue. We may be helped or hindered in cutting through the semantic difficulties of learning, training, functions, etc. in order to eventually arrive at a general theory of performance or behavior.

Bases of Classification

The next portion of this review proceeds from a consideration of the purposes of representative behavioral classifications which have been proposed, to an examination of the bases of behavioral taxonomy, insofar as they have been stated or can be inferred. Since the subject matter to be classified, in the systems we have studied, has generally been the "task," our concern must focus on two issues. First, is there consistency in the definition and meaning of the concept "task"? Second, is there general agreement as to the bases upon which task description, differentiation, and classification can be accomplished?

Definition of the Task

The concept of task has been defined in an almost endless variety of ways. In considering the general problem of task definition, there are two dimensions in terms of which major distinctions among

alternative definitions become apparent. The first of these is the scope or breadth of definition. Second is the extent to which tasks are conceptualized as external to and imposed upon subjects in contradistinction to the subject's interpretation of the nature of the task.

Task definitions vary greatly with respect to their breadth of coverage. At one end of this dimension are definitions which view the task as an integral part of and indistinguishable from a larger and more general work situation to which the individual is exposed. In this context the task is the utter totality of the situation imposed upon the subject. For example, this definition would consider ambient stimuli as an integral part of the task. The other end of this dimension is represented by definitions which treat a task as a specific <u>performance</u>. In this case, for example, one task could be to "depress the button whenever the light comes on." Suffice it to say that very different concepts may underlie definitions falling at either end of this dimension.

This diversity of opinion is also reflected in the extent to which tasks are defined as being external to or an intrinsic part of the subject. Some definitions take into account the propensity of subjects to redefine an imposed task in terms of their own needs, values, experiences, etc. In the grossest sense, these definitions treat a task as whatever it is the subject <u>perceives</u> the task to be. To that extent, the task is ideosyncratically and subjectively defined. Other definitions attempt to bypass the redefinition problem. They define the task in terms of what has been <u>imposed</u> upon the subject, be it total situation or a specific performance requirement.

Applying these dimensions to the task classification studies which we have reviewed is difficult. Although the concept of task was fundamental to most of the systems which were sampled, there were few rigorous attempts to define precisely what was meant by the concept.

In general, most investigators seemed to treat tasks as dynamic entities consisting of interrelated processes and activities. For example, R. B. Miller (1966) states that, "A task is any set of activities, occurring at about the same time, sharing some common purpose that is recognized by the task performer" (p. 11). Teichner and Olson (1968) share this dynamic view but in a more specific sense. They define a task, "...as a transfer of information between (system) components" (p. 4).

There is, however, another type of task definition which is employed by relatively few investigators. This type of definition imbues the concept of task with an objective existence clearly apart from the activities or processes which the operator subsequently brings into play. In this sense, the task is a set of conditions which elicits specific activities or processes. Hackman (1968) represents this point of view in his definition of a task:

> "A task is assigned to a person (or group) by an external agent or is self-generated, and consists of a stimulus complex and a set of instructions which specify what is to be done vis-a-vis the stimuli. The instructions indicate what operations are to be performed by the performer(s) with respect to the stimuli and/or what goal is to be achieved" (p. 12).

The extent to which these or any other definitions are "appropriate" can only be seen in the implications which they have for the problem of classification. We should not attempt to debate about <u>the</u> definition of a "task" as if only one were possible. Rather, we must adopt or <u>develop</u> a definition which will serve as an adequate vehicle for classification. An adequate vehicle will permit the derivation of terms which reliably describe tasks and distinguish among them. These derived terms provide the conceptual basis for classification.

Conceptual Bases for Classification

Every system of classification has as its very foundation a set of terms to be employed in the description and eventual classification of the subject-matter toward which the system is oriented. The investigators whom we have reviewed differ substantially in opinion as to the proper basis for describing and classifying tasks. Each has attempted to coin his own set of terms. And yet, differences of opinion on this issue are not unreasonable when one considers the diversity of objectives and the range of task definitions associated with these same investigators. Nevertheless, these alternative bases must be examined carefully. It must be determined whether they merely reflect preferences or whether they can be differentiated on more substantive grounds.

To lend structure to the appraisal, we shall deal with four, major, conceptual bases underlying task description and classification as discussed by Hackman (1968). The first three of these conceptualizations focus on the description of the operator's behaviors and abilities. Quite naturally, descriptors of this type accompany task definitions which are process, activity, or even more candidly, performance oriented. The fourth and final conceptualization of appropriate descriptors stems from a task definition which emphasizes the elements, conditions, or components of the task. These constituent elements are thought of as variables which may be manipulated to call different packages of abilities into play, to emphasize certain processes over others, in short, to affect performance.

Behavior Description Approach. In this conceptual approach to task classification, categories or types of tasks are to be formulated based upon observations and descriptions of what operators actually do

---

* This label and the three which follow were originally suggested by Altman (1966) and McGrath and Altman (1966).

while performing a task. Emphasis is placed primarily upon a description of the operator's overt behaviors in response to the task rather than on an analysis of what he is required or expected to do in order to produce criterion levels of performance. The data collected by time and motion analysts and by individuals engaged in task analysis are representative of the inputs upon which this approach is based.

Obviously, many variations on this theme are possible. While overt behaviors such as dial setting, meter reading, soldering, etc. are most often employed, certain subjective (and primarily cognitive) terms are also permissible (e.g., analyzing, computing, decision making, etc.). Variations are possible not only because of the sheer number of such terms, but also because many levels of description are possible (e.g., adjusts volume control, adjusts control, adjusts; depresses, depresses keys, types, etc.).

In spite of the large number of terms available for this approach to task description, relatively few descriptive systems have been developed which are based exclusively on operator behaviors or activities. Berliner, et al. (1964), for example, developed a hierarchical descriptive system, only the lowest or most specific level of which was oriented toward actual behaviors. In a more extensive development of this type of descriptive system, Reed (1967) has constructed a list of verbs which represent frequently encountered behaviors or activities occurring during task performance. Reed's list of activities is particularly impressive because of his attempt to minimize redundancy among terms and yet be fairly exhaustive in his coverage.

McCormick (1964, 1965, 1968) has employed this descriptive approach in his studies of worker-oriented job variables. He has dealt with the specification of human behaviors (e.g., handling objects, personal

contact with customers), as opposed to the more technological aspects of jobs, for the purpose of establishing common denominators across jobs. This effort has led to the development of his Position Analysis Questionnaire (PAQ) which represents a reasonably objective job analysis technique. Each of 189 job elements in the PAQ is to be rated in terms of its relevance to the job. More recently, McCormick (1968) has attempted to specify the human attributes that are relevant to the kinds of activities or behaviors occurring in various jobs. This interest in human attributes represents a transition from the behavioral description approach to the ability requirements approach discussed later in this paper.

S. A. Fine, beginning with his association with the U. S. Department of Labor, has also been interested in describing jobs on the basis of worker functions or behaviors. Terms of interest to him have included handling (things), analyzing (data), negotiating (with people), etc. Working with these and similar concepts, Fine (1963) is attempting a broad mapping of work behaviors along lines analogous to McCormick's job description efforts.

The extent to which the behavioral descriptive approach employed by these and other investigators could serve as a basis for task classification is questionable. In any relatively complex task myriad activities may be observed. Are descriptions of each component activity necessary to completely identify the task and reliably distinguish it from others? If each and every activity is not to be included how are the most "critical" or "representative" activities to be chosen--on the basis of frequency of occurrence, duration, criticality, etc.? As noted above, the system has been most often employed in the description of what people do on jobs.

Toward what ends would this type of system be most powerful? Hackman (1968) doubts whether it would be useful in "understanding how tasks affect behavior." He continues (p. 7):

> "It appears that some researchers concerned with job and task descriptions have, in effect, substituted a dependent variable class for what should be an independent variable class. That is, if we are interested in the effects of tasks and task characteristics on behavior, it is essential that we develop means of describing and classifying our independent variables (tasks) other than in terms of the dependent variables (behaviors) to which we ultimately wish to predict."

With other purposes in mind, such as building an information retrieval system which catalogues the effects of selected environmental variables on specific types of behavior, this approach might be of use. The point seems to be that acceptance or rejection of this (or any other) approach may only be possible in light of one's purpose in classifying. The issue is critical and will be raised again when we attempt to summarize this section.

Behavior Requirements Approach. There is a second approach to "task" description which appears to be increasing in popularity. In this approach emphasis is placed on the cataloging of behaviors which should be emitted or which are assumed to be required in order to achieve criterion levels of performance. The human operator is assumed to be in possession of a large repetoire of processes which will serve to intervene between stimulus events and output responses. Particular input-output configurations require that certain intervening processes or functions be called into play. Steiner (1966) summarizes this position in the following manner:

> "Task demands specify the kinds and amounts of resources that are needed, and the utilization pattern that is required if maximum productivity is to be obtained."

Many of the classificatory or descriptive systems which have been reviewed are based on the behavior requirements approach. There has been a great deal of interest in codifying the required intervening processes (functions, behaviors, etc.), cataloging tasks in terms of the types of processes required in successful performance, and then relating the types of tasks which emerge to particular training methodologies.

Although Gagne (1962), Gagne and Bolles (1963), and R. B. Miller (1966) are perhaps most representative of this approach, others have made use of it. For example, Eckstrand (1964), Folley (1964a), Annett and Duncan (1967), and E. E. Miller (1959) have all discussed or proposed the classification of tasks in terms of required behaviors. Considered collectively, these and related papers hypothesize initial lists of the major types of behavioral processes required in task performance, consider techniques for their detection or identification, and specify additional factors (sequencing of behaviors, time constraints, etc.) which should be considered for complete description of tasks in behavioral terms.

R. B. Miller (1966) provides a list of terms typical of the types of processes or functions used to describe and differentiate among tasks. These include: a scanning function, identification of relevant cues function, interpretation of cues, short-term memory, long-term memory, decision making and problem solving, and an effector response. Many of these terms are undoubtedly familiar to the reader. Familiar also may be the difficulties involved in analyzing tasks on the basis of these or similar terms. As Miller himself suggests, "The definitions --

even in their more extended and refined form -- are ambiguous in observing activities. They lack handles for quantification" (p. 13). In spite of these difficulties, however, Miller feels that this approach has utility, particularly for procedure design and development of appropriate training sequences. However, whether or not this system of description actually proves to be useful is as much a function of the "knowledge and skill" of its users as it is of the system itself. That the investigators who have developed such systems can employ them effectively is not enough. They must be made public, being useful to others as well as to the originators of the systems.

Ability Requirements Approach. The third conceptual basis for the description and classification of tasks is in many respects similar to the behavioral requirements concept. Tasks are to be described, contrasted and compared in terms of the abilities which a given task requires of the operator. These abilities are relatively enduring traits or attributes of the individual performing the task. The assumption is made that specific tasks will require certain ability profiles if performance is to be maximized with respect to established criteria. Tasks requiring similar ability configurations (both in terms of type and amount of ability) would be placed within the same category or said to be similar.

Fleishman (1967), Guilford (1967) and Thurstone (1944) have provided lists of abilities within the perceptual-motor, physical proficiency, cognitive, and perceptual domains. The ability requirements approach would require the analysis of tasks to determine the contributions to performance of these abilities or similar personal factors. Fleishman (1967) suggests that the various abilities can be thought of as "representing empirically derived patterns of response consistencies to task requirements varied in systematic ways" (p. 352). As mentioned

earlier, he has found this framework useful in the study of skill learning, training, component-total task relationships, prediction of retention, etc.

The abilities approach differs from the behavior requirements approach primarily in terms of concept derivation and level of description. The ability concepts are empirically derived through factor-analytic studies, and are treated as more basic units than the functions and processes posited by others. These differences not withstanding, the ability and behavior requirements approaches to classification proceed in the same manner. Both seek to identify critical aspects of the individual intervening between input stimuli and output response.

Because they proceed in a similar manner, they also have many problems in common. Chief among these is the subjective manner in which abilities (and processes or functions) are semantically defined. The abilities (i.e., the factors on the basis of which the abilities are inferred) are empirically derived from patterns of response consistencies on different tasks. The empirical definition for each factor is mathematically derived and consists of the factor-loadings of each of the many tasks contributing to the factor. This type of empirical-mathematical definition is unsatisfactory, however, when the factor-analyst wants to discuss his factors with other investigators, particularly in terms of distinguishing among factors I, II, III, etc. As a consequence, the attempt is invariably made to translate the empirical-mathematical definition of each factor into a semantic definition. This is accomplished by a "labeling" process.

The investigator carefully examines each task associated with a particular factor and gradually develops a set of hypotheses as to what it is that tasks loading on a given factor have in common. Even though he may find they have many things in common, he will inevitably try

to summarize them by employing a single label (e. g., perceptual speed, flexibility of closure, control precision). He will then define his label semantically in terms of the attributes which he feels best represent the communality among tasks on a specific factor. For example, Fleishman (1964) provides a semantic definition of the factor labeled "control precision."

> "This factor is common to tasks which require fine, highly controlled, but not over-controlled muscular adjustments, primarily where large muscle groups are involved... This ability extends to arm-hand as well as to leg movements. It is highly important in the operation of equipment where careful positioning of controls by the hands or feet are required. It is most critical where such adjustments must be rapid but precise..." (p. 16).

There is nothing capricious about this or any other ability definition. It was painstakingly developed from a rational analysis of tasks. The point remains, however, that even trained judges might have difficulty in analyzing tasks with respect to the abilities which are required. In order to make such judgments, the task analyst must have an idea about the <u>properties of the task itself</u> which place demands upon different abilities. This information is available but tends to be rather private, residing within the investigator who originally supplied the ability labels. It may, therefore, be possible to increase the descriptive power of this approach by asking the ability theorists to make the underlying communalities public.

<u>Task Characteristics Approach</u>. The fourth and final approach differs from the preceding approaches in terms of the type of task description which is attempted. The three approaches previously discussed are predicated upon task definitions which are process,

function, behavior, or performance oriented. Consequently, appropriate descriptive terms are those which focus on the task performer's overt activities or internal processing. Different tasks will evoke different activities, will require different types and sequences of processing, and will place demands on various configurations of abilities. In other words, task description focuses on what transpires between input and eventual output.

The task characteristics descriptive approach is based upon a rather different conceptualization of the "task." The approach is predicated upon a definition that treats the task as a set of conditions which elicit performance. These conditions are imposed upon the operator and have an objective existence quite apart from the activities they may trigger, the processes they may call into play, or the abilities they may require of the operator. Having adopted this point of view, appropriate descriptive terms are those which focus on the task _per se_. The assumption is made that tasks can be described and differentiated in terms of intrinsic objective properties which they may possess. These properties or characteristics may pertain to the goal toward which the operator works, relevant task stimuli, instructions, procedures, or even to characteristics of the response(s) or the task content. The obvious problem is the selection of those task components which are to be described as well as the particular terms or parameters by means of which description is to be accomplished. This selection will be dictated primarily by the particular task definition which is adopted.

Although several investigators (e.g., Hackman, 1968; Arnoult, 1963; Sells, 1963) have considered task classification on the basis of objective task properties, it has also been argued that the approach is unfeasible. This argument is founded on the problem of selecting some manageable set of descriptors. From the fantastic number of descriptors which are most certainly available, how does one decide upon which terms

to use? Lacking criteria for choices about types or levels of description, is there any alternative to the use of all possible terms? If not, we may be dealing with an unfeasible approach because of the staggering amount of work involved.* An extension of this argument suggests that the approach would eventually dissolve into S-R reductionism permitting study of the trees (S-R relationships) but providing little information about the forest (task performance).

Studies by Cotterman (1959), Fitts (1962), and Stolurow (1964) are perhaps most representative of attempts to pursue the task characteristics approach. They explicitly considered prototype task classifications based, at least in part, upon the description of task properties *per se*. In addition to an interest in the processes or functions evoked by a task, they were also concerned with morphological description. Stolurow (1964) succeeded in developing an extremely limited prototype classification of paired associate and serial learning tasks. Based upon his system of classification he was able to explore the effects of certain principles of learning (e.g., massed and distributed practice) with respect to different categories of tasks. Preliminary data suggested that the effects of massed and spaced practice were relatively homogeneous within certain of his task categories. This is especially interesting when one considers the types of descriptors upon which he based task categories (e.g., number and sequence of "stimuli" and responses; "stimulus" and response limits; meaningfulness; and qualitative relationships between "stimulus" and response).

---

* A similar problem has faced biological taxonomists. One school of thought, numerical taxonomy, argues for the inclusion of all possible characteristics. For example, Michner and Sokal (1957) employed 11,834 characteristics to differentiate among 97 species of bees. They admit that not even this set of terms was completely exhaustive!

Summary of Conceptual Bases of Classification

In this section, the general problems of task definition and selection of conceptual bases for task classification have been described. Satisfactory resolution of these issues is crucial for progress on the general classification problem. We have seen that tasks can be defined in several ways, particularly in regard to: the scope of definition; the extent to which tasks may be treated as objective entities, clearly apart from the operators who perform them; and the extent to which tasks are viewed as processes or structures.

The basic point is that there are few guidelines for selecting a particular definition or a particular approach to classification. If one is interested in how tasks and task properties affect behavior, then one might agree with Hackman (1968) that the task characteristics approach is reasonable. If one's purpose is to organize information on tasks, performance, and the variables affecting performance, the problem may be only slightly different. Such an information retrieval system must permit consistent and reliable comparisons among whatever indexing terms are employed (i.e., task properties, activities, functions, or abilities).

Of one thing we are certain. Sophisticated classification systems in vogue twenty years from now will, to some extent, incorporate many different conceptual systems. Among these, physiological descriptive systems surely will receive a big play. But our real dilemma is more immediate. Upon what conceptual base should we initially construct our system? Which conceptual foundation will promote and insure construction of a solid and habitable structure?

If behavioral classification is to provide information not only on what affects performance, but also why such effects occur, then an

approach is available. As suggested in the recommendations section of this report, the task characteristics approach, when coupled with specifications of the operator, may provide a convenient point of departure.

Procedures for Classification

Having briefly treated the taxonomic issues of purpose and conceptual basis, we are now in a position to consider the problem of "how" to proceed with classification. That is, how are tasks (no matter how defined) to be systematically arranged into groups or categories? Few behavioral scientists have discussed alternative procedures for classification or have spelled these out in any detail. The tendency has been to treat procedures in only the most general manner, discussing them perhaps, but seldom actually employing them.

Three issues arise in attempting to actually develop a classification system, given that an objective has been stated and that a conceptual basis for description has been chosen. The first of these issues is self-evident. It is mandatory that the subject matter (tasks) be classified as reliably as possible. The extent to which tasks can be classified reliably will depend upon the ability to objectively and operationally define task descriptors. The second issue is less obvious. Classification may proceed on qualitative or quantitative grounds. In other words, classes may represent different kinds of things (e. g., elephants or bananas, tracking tasks or decision-making tasks) or they may represent differences in degree with respect to dimensions which they have in common (e. g., millipedes, centipedes, spiders, quadrapeds, etc., classified with respect to number of feet). The third and final issue underlying classificatory procedures involves the selection and use of criteria to assess the adequacy and utility of classification. Each of these issues shall be treated briefly.

Operational Definitions

Regardless of the descriptors (unit characteristics) in terms of which classification is to be achieved, they ideally should be defined in operational terms permitting some form of quantitative assessment. That is, they should be reasonably objective. Kimble (1964) dwells on this issue and points out that the reliability with which distinctions among attributes can be made is largely a function of the extent to which they have been operationally defined. For example, Stolurow (1964) presented a series of task descriptions to a sample of judges. They were to analyze the tasks with respect to a list of quasi-operationally defined characteristics (e.g., number of stimuli, stimulus sequence, meaningfulness, etc.). Although inter-judge agreement was not reported, it may have been less than adequate, for Stolurow wished to more carefully redefine his descriptors. We have encountered similar difficulties in attempting to analyze tasks in terms of ability requirements. These examples (and many others which might have been picked) stress the need to develop objective and concise definitions which will permit clear and consistent distinctions among descriptors initially and among categories of tasks ultimately. A step in this direction has been taken by Teichner and Olson (1968) who have attempted to make their terms objective and public. They have identified four basic types of tasks (i.e., searching, switching, coding, and tracking). Each type of task is defined semantically and then operationally in terms of the type of response measure associated with it (i.e., probability of detection, reaction time, percent correct responses, and time on target).

The requirement for concise and objective definition of terms becomes increasingly critical if the system is to be used by a broad range of specialists. The descriptive terms may be completely unfamiliar to many of these individuals, or even too familiar as in the case of the popular terms "decision-making" and "problem-solving." As a consequence,

most taxonomists provide for the orientation or training of system users in the meaning of nomenclature. This, of course, should always be done. But the effectiveness of this training will only be as good as the definitions being learned. Objectivity of definitions bears directly on the reliability and precision with which the selected subject matter can be classified. It is critical to both qualitative and quantitative classification systems.

Qualitative and Quantitative Classification

The taxonomic issues of purpose, conceptual bases, and procedures are inextricably interwoven. Each step in the development of a classification system has implications for subsequent (and often preceding) steps. This interdependency is particularly noticeable when one attempts to specify the functional and structural nature of his system. The purpose for which classification is being attempted has implications for the definition of tasks. Task definitions directly determine the appropriate conceptual bases for task description and differentiation. The degree to which classificatory descriptors can be operationally and objectively defined appears to have implications for the system of measurement to be employed during classification. Similarly, the system of measurement which is adopted may well dictate the structural and functional nature of the classification system. This section focuses on the two latter aspects. That is, what systems of measurement are typically applied to descriptors, and what implications do they have for the structural and functional characteristics of the classification system?

Qualitative Classification.

As a minimum requirement, the descriptors employed in the differentiation and classification of tasks must permit nominal scaling. That is, a judge must at least be able to ascertain whether each descriptor applies or does not apply to the particular task being examined. Is the descriptor "present" or "absent"? In essence, the judge is required to make a series of qualitative

judgments about the presence or absence of a set of descriptors which are treated as attributes. Classification based upon this type of qualitative analysis can proceed in two ways. Biologists (Sokal & Sneath, 1963) describe these approaches as monothetic or polythetic classification.

In monothetic classification, the taxonomist defines each type of task or each category in terms of an unique and usually small set of attributes such that possession of these features is both necessary and sufficient for membership in the group so defined. The groups which result are termed monothetic groups because each grouping has an unique set of defining attributes. The taxonomist who employs this method of classification is essentially providing a priori differential weightings to a large set of attributes. Dealing with the set of attributes "A" through "Z", he may define a "monitoring task" as a task in which attributes "A", "B", and "C" are present. Similarly, he may define a "tracking task" as one in which attributes "X", "Y", and "Z" are judged present. In other words, he arbitrarily assigns a weight of one to those attributes which define a particular type of task and a weight of zero to those attributes not included in the definition of that same type of task. Finding attributes "X", "Y", and "Z" present in a specific task (e.g., aiming a visually sighted anti-aircraft weapon) he categorizes the task -- "it's a tracking task." This is essentially how Teichner and Olson (1968), for instance, might decide whether a particular specimen is a searching, switching, coding, or tracking task. If the response measure involves reaction time, then it is a switching task.

Other investigators whom we have reviewed would proceed in a slightly different manner. They attempt to develop polythetic classifications from an examination of the overall pattern of features which can be attributed to each specimen. Tasks which have the greatest number of shared features (e.g., the attributes "A" through "Z") would be placed

within the same category. No single attribute is either essential to group membership or is sufficient to make a particular task a member of the group. For instance, taxonomists who pursue this approach contend that any particular task will require some set or pattern of abilities from the total set of abilities, or some set or sequence of processes from a larger set. Classification proceeds as if a check-list were being employed. Figure 1 illustrates this polythetic approach. Each task is examined in terms of the presence (+) or absence (-) of each descriptor or attribute in the total set. The resultant pattern or configuration of pluses and zeros is used to describe and classify tasks. Those tasks having identical or "similar" patterns of attributes are placed within the same category.

	Tasks to be Classified				
Descriptors	1	2	3	4	m
1	+	+	0	+	.
2	+	0	0	+	.
3	+	+	+	0	.
4	+	0	+	0	.
.	.	.	.	.	.
.	.	.	.	.	.
n	.	.	.	.	.

Figure 1. Qualitative Classification Based on Attributes

An example of polythetic qualitative classification based on a nominal scale of measurement is one entertained by R. B. Miller (1966). As presented in Figure 2, Miller considers a prototype taxonomic grid consisting of three axes. The specific descriptors accompanying each axis are intended to be illustrative only. The important point is that judges would be required to examine tasks with respect to each axis of

description. Attributes on each axis would be judged present or absent,* yielding an unique configuration. Tasks having identical or "similar" configurations would be said to belong to the same class.

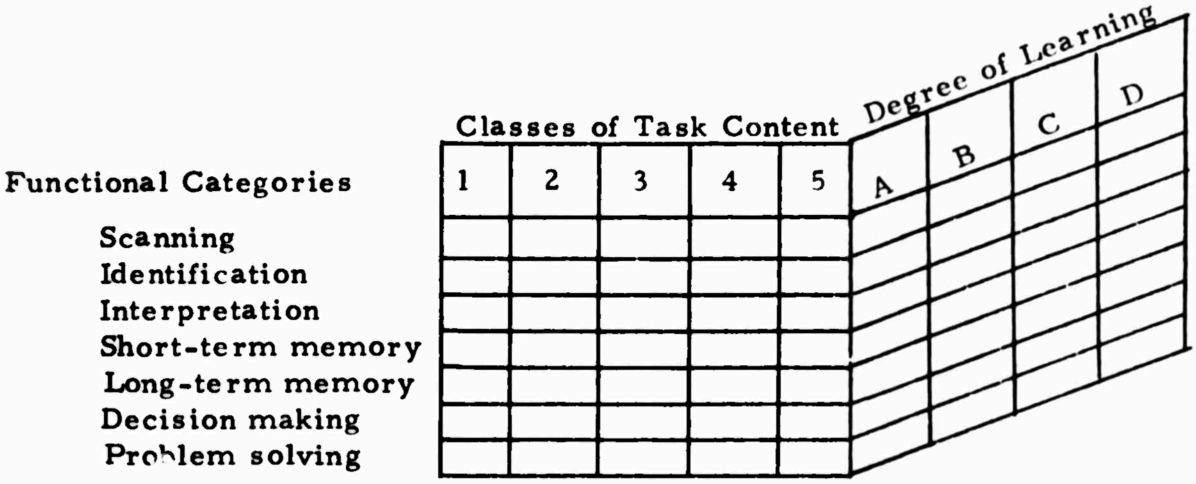

Figure 2. A Taxonomic Grid

Regardless of which particular approach (monothetic or polythetic) is chosen, classification on qualitative grounds is clearly possible given a set of reasonably well-defined descriptors whose presence or absence can be reliably determined for any particular task. Indeed, most of the investigators whom we have reviewed treat the process of classification as the categorization or "pigeon-holing" of tasks based on the pattern of attributes which they are judged to possess. Because the general qualitative approach seems to enjoy such popularity**, it is

* As an interesting dilemma, consider the qualitative assessment of two tasks in terms of 100 descriptors. Ninety descriptors are judged "absent" in both samples. The remaining ten descriptors are judged "present" in one sample but not in the other. Can we conclude that the two samples are highly similar? There appears to be a breach of logic in concluding that the fewer attributes either sample possesses, the more similar they are. Proof by exclusion, after all, has its limits.

** Few investigators have actually attempted classification. Consequently, one must be careful in discussing systems still at a conceptual, or at best, prototypic level of development. Most investigators seem to be interested in qualitative classification initially. We can only assume that their systems would become more quantitative during later stages of development.

absolutely mandatory that its implications for the classificatory structure be clearly understood. Classifications of this type have one characteristic property. The relationships among classes or categories of tasks are indeterminate.

To simplify this discussion, let us assume that we have a descriptive system based on four attributes -- A, B, C, and D. In assessing the presence or absence of these descriptors in a sample of tasks, the configurations shown in Figure 3 could be obtained.

	Sampled Tasks					
Descriptors	1	2	3	4	5	6
A	+	+	+	+	+	0
B	0	0	+	+	+	0
C	+	+	+	0	+	0
D	0	0	0	+	+	0

Figure 3. Qualitative Task Description

The important feature of Figure 3 is that there are two tasks which are <u>identical.</u> Based upon the four attributes which were employed, tasks # 1 and # 2 have precisely the same configurations. They belong to the same category or class by virtue of the presence of attributes A and C, and the absence of attributes B and D. But what can be inferred from the remaining tasks and the configurations of attributes associated with them?

At one level we could conclude that due to the vagaries of sampling, the remaining four tasks represent unique types or categories. With repeated sampling, other tasks might be found which exhibit one of these remaining four configurations. The extention of this argument is that there is a class for each possible configuration of attributes. For "n" attributes there are 2^n possible configurations. As "n" becomes large,

the number of possible categories would become extremely large. Ultimately, such a system would generate classes permitting extremely fine distinctions among tasks. In the extreme case, the distinctions would become so precise as to provide classes consisting of single tasks!

Consequently, the taxonomist must develop his classificatory system to get around the reductio ad absurdum problem. His approach is one of placing similar tasks within the same category. However, when one deals with the types of data shown in Figure 3, the determination of task similarities becomes extremely complex. What appears to be a fairly simple process can turn into an extremely frustrating experience. The particular criterion of similarity which is chosen is an arbitrary matter. There is no compelling logic for the adoption of one criterion over others. Consequently, for the data shown in Figure 3, a number of alternative groupings are possible depending on how one chooses to define "similarity." Perhaps the only point on which there would be agreement is that task # 6* is similar to no other configuration.

The considerations above are obviously important, but they are really tangential to the major problem in qualitative classification systems to which we alluded earlier. No matter how classes are generated, the relationships among them cannot be determined. In other words, the similarity among classes cannot be established. Dealing with nominal data, distance functions cannot be employed to express the degree of similarity between classes. Therefore, we are unable to determine whether, for example, class # 1 is more similar to class # 3 than it is to class # 5 because we lack the dimensions along which

---

* The configuration accompanying this task immediately suggests that additional attributes be employed. One gains little descriptive or classificatory information when he finds a sample which possesses none of the attributes included within the system of description.

such comparisons can be made.

Implied is the notion that classification systems based on nominal data will contain categories which differ in kind but not in degree. Simple though this idea may seem, it has important implications with respect to one's ultimate purpose in developing a classification system. If one's interest is in indexing the experimental literature, then categories differing in kind may suffice. However, if one's interest is in manipulating a task parameter and predicting the effect of such a manipulation upon performance, then task categories which differ in degree will not only be helpful but necessary. Such categories may be obtained by employing a quantitative approach to classification.

Quantitative Classification. Although classifications may be readily developed from qualitative differences among tasks, they might be based alternatively on quantitative differences. Yet, most behavioral taxonomists do not appear to have entertained classifications based explicitly upon the quantitative scaling of task descriptors. Because of the rudimentary stages of development so typical of many of their systems, it would be premature to conclude that interest in quantification was completely lacking. With sufficient rigor in the definition of descriptive terms, judges could be asked to rate or scale each descriptor in terms of its involvement in a particular task. Were fully operational definitions available, measurement might proceed in terms of counts (number of controls) or in terms of quantitative dimensions (level of illumination).

If one were to pursue development of a quantitative system of classification he would undoubtedly try to select a set of descriptive dimensions common to or applicable in a wide variety of tasks. Were he successful in this regard, then all task specimens might be described

in terms of each dimension. As shown in Figure 4, the use of common dimensions would permit a profile of dimensional values to be developed for each task under investigation. More importantly, since all tasks would be evaluated with respect to all dimensions, tasks would be distinguished solely on the basis of degree.

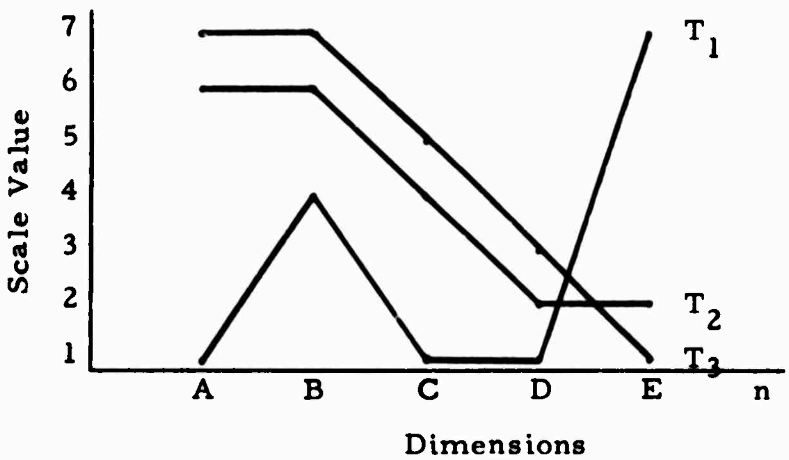

Figure 4. Task Profile Data

The problem of classification is one of determining the degree of similarity between task dimensional profiles. Those profiles having the greatest similarity would be placed within the same class. Silverman (1967) recommends that the numerical taxonomic procedures of the biologists be employed for this purpose. Although Silverman was primarily interested in employing such procedures in the development of occupational classifications, his suggested analytical approach is germane to the generation of classes of tasks. The numerical taxonomic procedures base task comparison and classification on operational and quantitative grounds. These techniques provide for the precise measurement of the similarity (distance) among the samples to be classified. This is generally accomplished by multidimensional scaling, cluster analytical, or discriminant function techniques. Each sample is located in hyperspace as a function of its values on the set of dimensions used

to describe it. Those samples which are located in the same general space tend to cluster or to fall into classes. The choice of a particular criterion for cluster size is arbitrary. However, once clusters are generated, the differences between them can be precisely described in terms of their distances from one another along each dimension of description. An oversimplified representation of classes developed in this manner is shown in Figure 5.

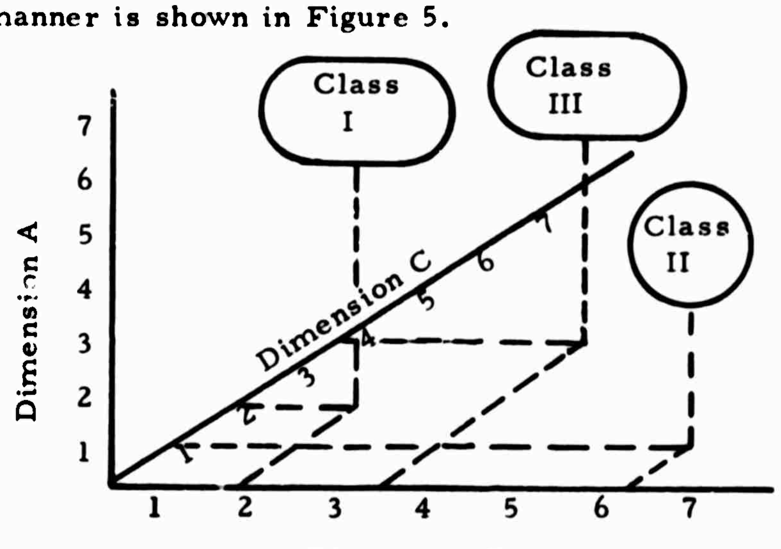

Figure 5. Numerical Classification

Numerical taxonomic techniques have not been applied yet to behavioral classifications. They should be fully explored, however, especially in light of the difficulties which behavioral taxonomists have in assessing the similarity among qualitatively derived categories. Indeed, without the use of such procedures it may be impossible to develop a classification within which hierarchical levels can be (quantitatively) established, and the relationships across and between levels fully specified.

Classificatory Criteria

Assume for the moment that the behavioral taxonomist has dealt successfully with the preceding issues. His purpose has been explicitly stated, he has chosen the subject matter to be classified, his descriptors have been carefully selected and defined. In short, he is ready to proceed to the matter of classification. At this juncture, however, an interesting problem arises. How may he judge the adequacy of the classification which is being developed? A number of criteria have been suggested for this evaluation.

First and foremost, descriptors must be defined and treated within a system of measurement so that they can be reliably evaluated. No matter what the descriptive bases or the techniques employed in classification, it is essential that descriptor values be assigned reliably. Reliability of description is the *sine qua non* of a reliable system of classification. Other criteria are typically called into play once the formal process of classification has been initiated.

One criterion frequently mentioned seems to have been inherited from the biological taxonomists. It requires that classes within the system be mutually exclusive. In other words, it should be possible to place any given sample of subject matter in one and only one classificatory grouping. If the classification contains a hierarchical structure, then categories on the same horizontal level are required to be mutually exclusive. Based upon our preceding discussion, it can be seen that this criterion will be most readily satisfied in monothetic qualitative systems, and hardest to achieve in quantitative systems.

Annett and Duncan (1967) suggest that behavioral taxonomists are not in complete agreement as to the necessity of mutual exclusivity. Some investigators apparently stress this criterion, attempting to

achieve it through careful formal definition of categories. Others seem to de-emphasize it, finding it to be unrealistically constraining at least during initial stages of classificatory development. Considered jointly these two stances assume a reasonable posture. One should probably strive for eventual exclusivity of classes, but he should also be willing to accept less rigor during his initial efforts.

A second major criterion concerns the extent to which classification of subject matter is exhaustive. Powerful classifications are viewed as those in which every sample of subject matter can be located. Annett and Duncan (1967) identify two reactions to this criterion by behavioral taxonomists. Some investigators have started initially by considering a tremendous range of samples with the objective of accounting for each in their systems. Others seemed to be initially interested in classifying a smaller set and provided a catch-all category for samples which could not be immediately located within their formal framework. Again, the suggestion is that while exhaustive classification is an ultimate objective, it is perhaps unrealistic during initial efforts. It is this continuing search for both mutually exclusive classes and exhaustive classification which seems to lead to the constant revision and modification found in healthy systems of classification.

A third major criterion is also of importance to the behavioral taxonomist. Its nature is self-evident. Eventually classes are desired which have specific behavioral implications. For example, Annett and Duncan (1967) are interested in classifying "tasks" so that each category or class of tasks has specific training requirements associated with it. Stolurow (1964) sought similar ties between task categories and specific principles of learning. Teichner and Olson (1968) are seeking classes of tasks for which selected environmental variables, such as level of O_2 saturation, have similar effects upon performance. Ultimately, of course, any behavioral classification scheme must make the "match" between specific

categories and behavioral effects. The degree to which the "match" can be made will determine the predictive power of the system. At one level a statement might concern whether or not a particular environmental variable would affect performance. At another and more sophisticated level it might be possible to predict the direction and magnitude of effect. We have attempted to show how quantitative systems, in particular, would meet this criterion.

A final set of criteria should also be mentioned. These are primarily practical, but are really no less important than those previously outlined. R. B. Miller (1966) suggests that the emergent classification be considered in terms of its efficiency and utility. It should promote communication among its users, be they laboratory researchers or technologists. It should enable them to meet their responsibilities more effectively. Miller cogently argues that perhaps the final and ultimate criterion is the degree of acceptance which the schema comes to enjoy. Effective handling of the preceding issues would certainly be a step toward such a goal.

Summary of Procedures for Classification

The general issue of "how" to classify has received relatively little explicit or detailed treatment in the literature. Investigators raise the general problems with which we have been concerned, but having raised them fail to achieve a satisfactory conclusion. This lack of critical thinking, on what is perhaps the most crucial taxonomic problem, undoubtedly comes from want of opportunities (in terms of time and money) to actually attempt classification. Indeed, it may only be through the actual development of provisional systems that these issues can be distilled and crystallized.

Homage is paid to the operational definition of unit characteristics, but few investigators have actually generated such definitions. Stolurow (1964) and Teichner and Olson (1968) have gone the furthest in the direction, while E. E. Miller (1966) and Haggard (1963) seem to be on the same track. In a similar vein, the literature provides few insights into the actual process of classification, particularly in regard to whether classification is to proceed along qualitative or quantitative paths. Stolurow (1964) initially employs a qualitative approach based upon the presence or absence of critical task characteristics. On the other hand, one investigator (Silverman, 1967) speculates on the applicability of quantitative, numerical taxonomic procedures. Other investigators provide no clear indication of how they would proceed.

More agreement exists in regard to major classificatory criteria. Ultimately desired is a system which permits exhaustive classification and which consists of mutually exclusive categories. It is the consensus of taxonomists that these two criteria be applied liberally during initial developmental efforts. Undue emphasis on these criteria during initial efforts is viewed as overly restrictive. Taxonomists agree completely on the third major criterion. Regardless of what is classified the system must eventually be tied to behavior or performance.

Future developmental efforts must meet these procedural questions head-on and solve them if a viable taxonomic system is to be produced. The implications of classifications based on differences in kind or in degree must be fully explored. This analysis must be conducted in light of concisely stated objectives and precisely defined unit characteristics. Perpetuating a general treatment of these topics will result in little of real value.

CONCLUSIONS AND RECOMMENDATIONS

The original purposes in preparing a report on behavioral taxonomy were twofold. On the one hand, it seemed reasonable to review relevant literature in order to assess the "state-of-the-art" of behavioral taxonomy and to seek guidelines for taxonomic efforts which might be conducted in the future. On the other hand, there was interest in applying a provisional set of criteria to each classification system encountered, so that the best available system might be focused upon, expanded, and refined. In retrospect these endeavors were only partially successful.

Although several alternative "systems" were examined with respect to procedures which might be adopted (or rejected) in future efforts, few specific guidelines were obtained. Instead, the review only served to raise, and to leave essentially unanswered, a number of taxonomic issues. These included the purpose in mind when classification is attempted, the definition of the task concept, the selection of descriptive bases, and the development of formal procedures for conduct of the classificatory exercise. Similarly, during early stages of the review an evaluation of alternative schemas was undertaken. As Ginsberg, et al. (1966) found in an earlier review, however, the rudimentary level of development so typical of these systems rendered systematic evaluation all but impossible. They simply have not been developed far enough for reasonable evaluative criteria to be brought into play.

Having reviewed the representative literature in this area, one is compelled to conclude that behavioral taxonomy is still in its infancy. In spite of ten years of thought and effort, relatively little has been accomplished. Primary emphasis seems to have been placed upon expressions of the need for behavioral classification systems. In the

relatively few instances in which this challenge has been accepted, conceptual schemas have outnumbered attempts to actually develop and apply systems of classification. Consequently, it is not surprising that there is little to go on, that few specific guidelines can be extracted from the literature, or that evaluations of alternative schemas cannot be readily undertaken.

Because of the lack of formal studies around which a general and informative review could be structured, the present paper has focused more on certain theoretical questions than was originally intended. By default, its purpose has become one of uncovering and presenting conundrums which may have impeded taxonomic progress in the past, and to which behavioral taxonomists might profitably address themselves in the future. The major issues involved in the development of "behavioral" or "task" classification systems are presented below.

1. Purpose - The first and prerequisite step in the development of a classification system is a precise and detailed statement of purpose which describes the specific context and manner in which the system would be applied. In light of the stated purpose one of two general systems will be more appropriate: a) utilitarian classification with specific applications; or b) theoretical classification with broad applications.

2. Subject-matter - Pains must be taken to identify and define the type of subject-matter which must be classified in order to achieve the stated objectives. If "tasks" are identified as the appropriate subject-matter, they may be defined generally as dynamic processes or as static structures.

3. Descriptive Bases - For the subject-matter definition which is adopted, appropriate descriptive terms must be generated which can

be used to compare and contrast samples of subject-matter. Differential description may be based upon: a) observed behaviors, b) required behaviors or processes, c) required abilities, or d) task characteristics.

4. Procedures of classification - The descriptive bases can be treated as sets of attributes or as dimensions. Depending upon the system of measurement which is employed, classes or categories can be generated by one of the following procedures: a) monothetic qualitative approach, b) polythetic qualitative approach, or c) quantitative approach.

The issues summarized above may be viewed as steps in the process of developing a system of classification. As uch they are interactive. Implementation of each step will depend on what has preceded it and will affect those steps which follow. These in turn may require reconsideration of the results of some earlier step. Consequently, development of a behavioral classification requires an iterative solution. Previous efforts have neither stressed nor capitalized upon these interrelationships.

The attempt has been made to present these issues objectively. In spite of this effort, the reader still may have detected certain biases. It is important, therefore, that they be made explicit. They stem from the author's involvement in the specific problem of developing a task classification system which will increase our ability to make predictions about facets of human performance. Consideration of this problem has convinced the author that the appropriate subject-matter to be classified is the "task", and that this concept can be most powerfully defined as a static structure representing a set of conditions imposed upon the task performer. This definition suggests that the appropriate basis for description and differentiation of tasks is the "task characteristics" approach. This approach emphasizes certain independent variables hypothesized to comprise tasks. This approach to description assumes that all tasks possess each of the critical variables but that the degree of possession varies from task to task. Finally, it

is felt that a powerful method of classification, for tasks described in this fashion, is one which quantitatively estimates the degree of similarity among tasks based upon the similarity among task characteristic profiles.

REFERENCES

Alexander, L., & Cooperband, A. S. A method for system task analysis using statistical decision theory. Santa Monica, California: SDC, August 1965.

Altman, I. Aspects of the criterion problem in small group research: II. The analysis of group tasks. Acta Psychologica, 1966, 25, 199-221.

Annett, J., & Duncan, K. D. Task analysis and training design. Occupational Psychology, 1967, 41, 211-221.

Arnoult, M. D. The specification of a "social" stimulus. In S. B. Sells (Ed.), Stimulus determinants of behavior. New York: Ronald Press, 1963.

Bergum, B. O. A taxonomic analysis of continuous performance. Perceptual and Motor Skills, 1966, 23, 47-54.

Berliner, D. C., Angell, D., & Shearer, J. Behaviors, measures, and instruments for performance evaluation in simulated environments. Paper delivered for a symposium and workshop on the quantification of human performance, Albuquerque, New Mexico, 1964.

Bloom, B. S. Taxonomy of educational objectives. The classification of educational goals. Handbook I: Cognitive domain. New York: David McKay Company, Inc., April 1967.

Cattell, R. B., & Warburton, F. W. Objective personality and motivation tests. Urbana: University of Illinois Press, 1967.

Cotterman, T. E. Task classification: An approach to partially ordering information on human learning. Ohio: WPAFB, January 1959. WADC TN 58-374.

Cotterman, T. E. Problems in describing and categorizing tasks for determining training needs. Paper given at Ohio Psychological Association, Dayton, Ohio, October 1959.

Eckstrand, G. A. Current status of the technology of training. Ohio: WPAFB, September 1964. AMRL-TR-64-86.

Fine, S. A. A functional approach to a broad scale map of work behaviors. McLean, Virginia: Human Sciences Research, Inc., September 1963. HSR-RM-63/2.

Fitts, P. M. Factors in complex skill training. In R. Glaser (Ed.), Training research and education. New York: University of Pittsburgh Press, 1962.

Fleishman, E. A., & Hempel, W. E., Jr. The relation between abilities and improvement with practice in a visual discrimination reaction task. Journal of Experimental Psychology, 1955, 49, 301-312.

Fleishman, E. A., & Hempel, W. E., Jr. Factorial analysis of complex psychomotor performance and related skills. Journal of Applied Psychology, 1956, 40, 96-104.

Fleishman, E. A., & Parker, J. F. Factors in the retention and relearning of perceptual-motor skill. Journal of Experimental Psychology, 1962, 64, 215-226.

Fleishman, E. A. The structure and measurement of physical fitness. Englewood Cliffs, New Jersey: Prentice-Hall, 1964.

Fleishman, E. A., & Fruchter, B. Component and total task relations at different stages of learning a complex tracking task. Perceptual and Motor Skills, 1965, 20, 1305-1311.

Fleishman, E. A. (Baker, W. J., Elkin, E. H., Van Cott, H. P., & Fleishman, E. A.) Effects of drugs on human performance: The development of analytical methods and tests of basic human abilities. Washington, D. C.: American Institutes for Research, 1966. AIR-E25-3/66-TR-2.

Fleishman, E. A. Development of a behavior taxonomy for describing human tasks: A correlational-experimental approach. Journal of Applied Psychology, 1967a, 51, 1-10.

Fleishman, E. A. Performance assessment based on an empirically derived task taxonomy. Human Factors, 1967b, 9, 349-366.

Folley, J. D., Jr. Development of an improved method of task analysis and beginnings of a theory of training. Port Washington, New York: USNTDC, June 1964a. NAVTRADEVCEN 1218-1.

Folley, J. D., Jr. Guidelines for task analysis. Port Washington, New York: USNTDC, June 1964b. NAVTRADEVCEN 1218-2.

Gagne, R. M. Human functions in systems. In R. M. Gagne (Ed.), Psychological principles in system development. New York: Holt, Rhinehart, & Winston, 1962.

Gagne, R. M., & Bolles, R. C. A review of factors in learning efficiency. In E. Galanter (Ed.), Automatic teaching: The state of the art. New York: John Wiley & Sons, Inc., 1963.

Ginsberg, Rose, McCullers, J. C., Meryman, J. J., Thomson, C. W., & Witte, R. S. A review of efforts to organize information about human learning, transfer, and retention. San Jose, California: San Jose State College, 1966.

Grether, W. F. Engineering psychology in the United States. American Psychologist, 1968, 23, p. 743-751.

Gilford, J. P. The nature of human intelligence. New York: McGraw-Hill, 1967.

Hackman, J. R. Tasks and task performance in research on stress. In J. E. McGrach (Ed.), Social and psychological factors on stress. New York: Holt, Rhinehart & Winston, 1968.

Haggard, D. F. The feasibility of developing a task classification structure for ordering training principles and training content. Fort Knox, Kentucky: Human Resources Research Office, 1963. Research Memorandum.

Kidd, J. S. Human tasks and equipment design. In R. Glaser (Ed.), Psychological principles in system development. New York: Holt, Rhinehart & Winston, 1962.

Kimble, G. A. Categories of learning and the problem of definition: Comments on Professor Grant's paper. In A. Melton (Ed.), Categories of human learning, New York: Academic Press, 1964.

McCormick, E. J. The development, analysis, and experimental application of worker-oriented job variables. Lafayette, Indiana: Purdue, 1964. Final Report; ONR Nonr-1100 (19).

McCormick, E. J. Job dimensions: Their nature and possible uses. Paper read at the American Psychological Association Convention, Chicago, 1965.

McCormick, E. J. Job dimensions: Their nature and possible uses Paper presented at the International Congress of Applied Psychology, Amsterdam, August 1968.

McGrath, J. E., & Altman, I. Small group research: A synthesis and critique of the field. New York: Holt, 1966.

Melton, A. W. The taxonomy of human learning: Overview. In A. Melton (Ed.), Categories of human learning. New York: Academic Press, 1964.

Melton, A. W., & Briggs, G. E. Engineering psychology. Annual Review of Psychology, 1960, 11, 71-98.

Michner, C. D., & Sokal, R. R. A quantitative approach to a problem in classification, Evolution, 1957, 11, 130-162.

Miller, E. E. Amplification of a system for classifying learning tasks. Urbana, Illinois: University of Illinois, 1959. Memorandum Report #3, AF Project: 33(616)-5965.

Miller, E. E. A taxonomy of response processes. Fort Knox, Kentucky: HumRRo, Division 2, December 1966. Study BR-8.

Miller, R. B. Task taxonomy: Science or technology? Poughkeepsie, New York: IBM, 1966.

Reed, L. E. Advances in the use of computers for handling human factors task data. Ohio: WPAFB, April 1967. AMRL-TR-67-16.

Sells, S. B. Dimensions of stimulus situations which account for behavior variance. In S. B. Sells (Ed.), Stimulus determinants of behavior. New York: Ronald Press, 1963.

Silverman, J. New techniques in task analysis. San Diego, California: U.S. Naval Personnel Research Activity, 1967. SRM 68-12.

Sokal, R. R., & Sneath, P. H. Principles of numerical taxonomy. San Francisco: W. H. Freeman & Co., 1963.

Steiner, I. D. Models for inferring relationships between group size and potential group productivity. Behavioral Science, 1966, 11, 273-283.

Stolurow, L. M. A taxonomy of learning task characteristics. Ohio: WPAFB, 1964. AMRL-TDR-64-2.

Teichner, W., & Olson, Diane. Predicting human performance in space environments. Boston: Harvard University, 1968.

Thurstone, L. L. A factor analysis study of perception. <u>Psychometric Monographs</u>, 1944, No. 4.

Wickens, D. D. The centrality of verbal learning: Comments on Professor Underwood's paper. In A. Melton (Ed.), <u>Categories of human learning</u>. New York: Academic Press, 1964.

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) American Institutes for Research 135 North Bellefield Avenue Pittsburgh, Pennsylvania 15213	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

DEVELOPMENT OF A TAXONOMY OF HUMAN PERFORMANCE: A REVIEW OF CLASSIFICATORY SYSTEMS RELATING TO TASKS AND PERFORMANCE

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Scientific Interim

5. AUTHOR(S) (First name, middle initial, last name)

George R. Wheaton

6. REPORT DATE December 1968	7a. TOTAL NO. OF PAGES 53	7b. NO. OF REFS 51

8a. CONTRACT OR GRANT NO. F44620-67-C-0116(ARPA) b. PROJECT NO. 9743 c. 61101D d. 681313	9a. ORIGINATOR'S REPORT NUMBER(S) AIR-726-12/68-TR-1 R68-7 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) AFOSR-69-1405-TR

10. DISTRIBUTION STATEMENT

1. This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES TECH, OTHER	12. SPONSORING MILITARY ACTIVITY Air Force Office of Scientific Research 1400 Wilson Boulevard (SRLB) Arlington, Virginia 22209

13. ABSTRACT

In recent years there has been a cresive interest in the science of taxonomy in general, and in behavioral classifications in particular. In this report, past efforts are reviewed in an attempt to assess the "state-of-the-art" and to provide procedural guidelines for future taxonomic efforts. Approaches to and dilemmas encountered in attempting to develop systems of classification are discussed. Special consideration is given to those systems oriented toward the organization and understanding of information about human task performance. Within this context the taxonomic issues of purpose, descriptive bases, and methodological approaches are discussed in terms of available alternatives. The report leads to the conclusion that behavioral taxonomy is still in its infancy and that truly powerful systems of classification have yet to be developed. The paper suggests how substantive progress may be made by attempting development of a task classification system based upon numerical taxonomic procedures.

DD FORM 1 NOV 65 1473

UNCLASSIFIED

Security Classification